



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/997,957	11/30/2001	Joachim Frank	DE920000055US1 (590.080)	5057
35195	7590	11/14/2006		EXAMINER
				VO, HUYEN X
			ART UNIT	PAPER NUMBER
			2626	

DATE MAILED: 11/14/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/997,957	FRANK ET AL.	
	Examiner	Art Unit	
	Huyen X. Vo	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on 18 July 2006.
 2a) This action is **FINAL**. 2b) This action is non-final.
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 19-49 is/are pending in the application.
 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
 5) Claim(s) _____ is/are allowed.
 6) Claim(s) 19-49 is/are rejected.
 7) Claim(s) _____ is/are objected to.
 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
 10) The drawing(s) filed on 30 November 2001 is/are: a) accepted or b) objected to by the Examiner.
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____ | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

Response to Arguments

1. Applicant's arguments filed 7/18/2006 have been fully considered but they are not persuasive. Bonastre fully anticipates the limitation regarding detecting a speaker change and recognizing speaker upon detection of a speaker change (*speaker verification/recognition is carried out in section 3*). Glickman et al. is relied upon for the teaching of the limitation regarding using different dictionary for each speaker in the speech recognition process (*col. 5, lines 30-67, transcribing speech using speech models of the identified speaker; each speech model related to the identified speaker is inherently associated with linguistic information or a word that produced the speech model. Thus, the speech models related to a particular user also include vocabulary words in association. And the vocabulary words are considered as a speaker-related dictionary*).

2. In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992).

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless – (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

4. Claim 47 is rejected under 35 U.S.C. 102(a) as being anticipated by Bonastre et al. (IEEE Publication).

5. Regarding claim 47, Bonastre et al. disclose a speech recognition processing an incoming audio stream containing human speech from a plurality of speakers and having at least two speaker models and/or speaker-specific dictionaries, comprising: a detector which detects a speaker change in the incoming audio stream (*sections 2.1-2.2 on page 1178 and referring to abstract section*); a gather which gathers speaker-specific information with corresponding speaker-specific information of at least one predetermined known speaker from among the plurality of speakers thus recognizing the at least one predetermined speaker (*sections 2-3.2, input speech is processed to extract speech features, which are then compared with speech models of each enrolled speaker to determine a match*); and an interchanger which interchanges between the at least two speaker-specific dictionaries dependent on the detected speaker change and the corresponding recognized speaker (*sections 2-3.2, extracted features must be compared with speech models of a plurality of speakers enrolled before runtime*).

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. Claims 19-20, 22-26, 28-31, 33-39, 41-46, and 48-49 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bonastre et al. (IEEE Publication) in view of Glickman et al. (US 6067059).

8. Regarding claims 19, 31, 34, and 48, Bonastre et al. disclose a method, apparatus, and a program storage device readable by machine for processing a continuous audio stream containing human speech from a plurality of speakers related to at least one particular transaction, comprising the steps of: identifying a known speaker from among the plurality of speakers (*abstract section page 117*); digitizing the continuous audio stream (*ADC is inherently included in a digital system*); detecting a speaker change in the digitized audio stream (*sections 2.1-2.2 on page 1178 and referring to abstract section*); performing a speaker recognition if a speaker change is detected (*section 3 on page 1179*); and wherein each speaker is processed using a different dictionary of different topics (*each enrolled speaker has their own models stored in the system before runtime*).

Bonastre et al. fail to disclose the step of transcribing at least part of the continuous audio stream if a predetermined speaker is recognized. However, Glickman et al. teach the step of transcribing at least part of the continuous audio stream if the known speaker is recognized (*col. 5, ln. 30-67*).

Since Bonastre et al. and Glickman et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Bonastre et al. by incorporating the teaching of Glickman et al. in order to provide automatic closed-caption using speaker-dependent models to enhance speech recognition accuracy.

9. Regarding claims 25, 35, 39, 43, and 49, Bonastre et al. disclose a method, apparatus, and program storage device readable by machine for processing a continuous audio stream containing human speech of a plurality of speakers related to at least one particular transaction, comprising the steps of: identifying a known speaker from among the plurality of speakers (*abstract section page 117*); digitizing the continuous audio stream (*ADC is inherently included in digital systems*); detecting a speaker change in the digitized audio stream (*sections 2.1-2.2 on page 1178 and referring to abstract section*); performing a speaker recognition if a speaker change is detected (*section 3 on page 1179*); and wherein each speaker is processed using a different dictionary of different topics (*each enrolled speaker has their own models stored in the system before runtime*).

Bonastre et al. fail to disclose the step of indexing the audio stream with respect to the detected speaker change if the known speaker is recognized. However, Glickman et al. teach the step of indexing the audio stream with respect to the detected speaker change if the known speaker is recognized (*col. 5, ln. 30-67, labeling “Bob” or “Alice” to transcribed text of corresponding audio segments*).

Since Bonastre et al. and Glickman et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Bonastre et al. by incorporating the teaching of Glickman et al. in order to enable the system to use speaker-specific speech recognition models for a particular speaker to improve speech recognition accuracy.

10. Regarding claim 42, Bonastre et al. disclose an apparatus according to claim 39, further comprising a monitor which continuously monitors a real-time continuous audio stream and performing the steps of: digitizing the continuous audio stream (*ADC is inherently included in a digital system*); detecting a speaker change in the digitized audio stream (*sections 2.1-2.2 on page 1178 and referring to abstract section*); performing a speaker recognition if a speaker change is detected (*section 3 on page 1179*). Bonastre et al. fail to disclose the step of transcribing at least part of the continuous audio stream if a predetermined speaker is recognized. However, Glickman et al. teach the step of transcribing at least part of the continuous audio stream if the known speaker is recognized (*col. 5, ln. 30-67*).

Since Bonastre et al. and Glickman et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Bonastre et al. by incorporating the teaching of Glickman et al. in order to provide automatic closed-caption using speaker-dependent models to enhance speech recognition accuracy.

11. Regarding claims 20, 26, 36-37, and 44-45, Bonastre et al. fail to disclose a method, apparatus and computer readable medium according to claims 19, 25, 31, and 39, comprising the further step of protocolling time information for detected speaker changes. However, Glickman et al. further teach the step of protocolling time information for detected speaker changes (*timing info 332 in figure 3*).

Since Bonastre et al. and Glickman et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to further modify Bonastre et al. by incorporating the teaching of Glickman et al. in order to improve alignment of audio segments with corresponding transcribed text segments.

12. Regarding claims 22-23, 28-29, 38, and 46, Bonastre et al. further to disclose a method, apparatus, and computer readable medium according to claims 19, 25, 31, and 39, wherein the step of detecting a speaker change is accomplished by use of at least one characteristic audio feature, in particular features derived from the spectrum of the audio signal (see *figure 2, parameter extraction and feature vector of speech signal*);

and wherein the step of performing a speaker recognition involves the particular steps of calculating a speaker signature from the audio stream and comparing the calculated speaker signature with at least one known speaker signature (see *figure 2, parameter extraction and feature vector of speech signal. Audio characteristics or speech features/parameters are signature of the target speaker*).

13. Regarding claims 24 and 30, Bonastre et al. fail to disclose a method and apparatus according to claims 19 and 25, for use in a speech recognition or voice control system comprising at least two speaker-specific speaker models and/or dictionaries, wherein interchanging between the at least two speaker-specific dictionaries dependent on the detected speaker change and the corresponding recognized speaker. However, Glickman et al. further teach a speech recognition or voice control system comprising at least two speaker-specific speaker models and/or dictionaries, wherein interchanging between the at least two speaker-specific dictionaries dependent on the detected speaker change and the corresponding recognized speaker (*col. 5, lines 43-62*).

Since Bonastre et al. and Glickman et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to further modify Bonastre et al. by incorporating the teaching of Glickman et al. in order to improve speech recognition accuracy.

Art Unit: 2626

14. Regarding claims 33 and 41, Bonastre et al. fail to specifically disclose an apparatus according to claims 31 and 39, further comprising a scanner which automatically scans a continuous audio record, in particular a continuous audio stream recorded on a data or a signal carrier, and for detecting speaker changes in the continuous audio record. However, Glickman et al. further inherently teach such a scanner (*col. 2, lines 23-37, audio and text data are stored as two files, and files are stored in conventional disks or memory*).

Since Bonastre et al. and Glickman et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to further modify Bonastre et al. by incorporating the teaching of Glickman et al. in order to enable the system to perform speaker change detection and recognition on any source of audio data.

15. Claims 21, 27, 32, and 40 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bonastre et al. (IEEE Publication) in view of Glickman et al. (US 6067059), as applied to claim 19, and further in view of Kimber et al. (US 5598507).

16. Regarding claims 21, 27, 32, and 40, the modified Bonastre et al. fail to disclose a method, apparatus, and computer readable medium according to claims 19, 25, 31, and 39, wherein the step of detecting a speaker change and/or the step of performing a speaker recognition is/are preceded by the further step of detecting non-speech boundaries between continuous speech segments. However, Kimber et al. further

teach wherein the step of detecting a speaker change and/or the step of performing a speaker recognition is/are preceded by the further step of detecting non-speech boundaries between continuous speech segments (*col. 12, ln. 1-10, specifically elements 212 or 216 in figure 12*).

Since the modified Bonastre et al. and Kimber et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to further modify Bonastre et al. by incorporating the teaching of Kimber et al. in order to improve speech recognition accuracy.

Conclusion

THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Huyen X. Vo whose telephone number is 571-272-7631. The examiner can normally be reached on M-F, 9-5:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

HXV

11/12/2006



RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER